

## ABSTRACT OF THE DISCLOSURE

The present invention provides a method and apparatus for scalable server clustering.

In one embodiment, a plurality of servers service client requests. In one embodiment,

- 5 responsibility for servicing a client is assigned to a primary server. Requests from the client are routed to the primary server. In one embodiment, the primary server is selected for a client using a round-robin method. In another embodiment, a server monitors how many clients it serves as a primary server. If a client is to be assigned a server as a primary server and the server determines that it is assigned to a sufficient number of clients, the server can
- 10 select a different server to serve as the client's primary server. In one embodiment, an alternate server is assigned to the client. If the primary server is unavailable, the alternate server becomes the primary server and a new alternate server is selected.

LA 44875v4 02/26/2019